Random Multi-Overlap Structures for Optimization problems

Luca De Sanctis *

February 2, 2008

Abstract

We extend to the K-SAT and p-XOR-SAT optimization problems the results recently achieved, by introducing the concept of Random Multi-Overlap Structure, for the Viana-Bray model of diluted mean field spin glass. More precisely one can prove a generalized bound and an extended variational principle for the free energy per site in the thermodynamic limit. Moreover a trial function implementing ultrametric breaking of replica symmetry is exhibited.

Key words and phrases: optimization problems, replica symmetry breaking, ultrametric overlap structures.

1 Introdution

In the case of non-diluted spin glasses, M. Aizenman R. Sims and S. L. Starr ([6]) introduced the idea of Random Overlap Structure (ROSt) to express in a very elegant manner the free energy of the model as the an infimum over a rich probability space, to exhibit an optimal structure (the so-called Boltzmann one), to write down a general trial function through which one can formulate various ansatz's for the free energy of the model. It was also described how to formulate in particular the Parisi ansatz within this formalism. In [4, 5] we extended those

*Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton NJ 08544–1000 USA <lde@math.princeton.edu>

results to the case of diluted spin glass (Viana-Bray model). Here we extend the same results to optimization problems, the K-SAT and the p-XOR-SAT. The latter is the simple extension to p-body interactions of the Viana-Bray model, which is the diluted version of the famous model of Sherrington and Kirkpatrick of mean field spin glass. Many of the calculations in the present paper are quite simple and standard, and as general reference with many details the reader can take for instance [1].

2 Model, Notations, Definitions

Consider configurations of Ising spins $\sigma : i \to \sigma_i = \pm 1, i = 1, ...N$. Let P_ζ be Poisson random variable of mean ζ, and $\{i_\nu^\mu\}$ be independent identically distributed random variables, uniformly distributed over points $\{1, \dots, N\}$. If $\{J_\nu^\mu\}$ are independent identically distributed copies of a symmetric random variable $J = \pm 1$, then the Hamiltonian of random K-SAT is

$$H = -\sum_{\nu=1}^{P_{\alpha N}} \frac{1}{2}(1 + J_\nu^1 \sigma_{i_\nu^1}) \cdots \frac{1}{2}(1 + J_\nu^K \sigma_{i_\nu^K}) \ .$$

Here $\alpha \geq 0$ is the degree of connectivity and both p and K are supposed to be even. We do not consider the presence of an external field, but all the results trivially extends to this case as well. By ω we mean the Bolztmann-Gibbs average

$$\omega(\mathcal{O}) = Z_N^{-1} \sum_{\{\sigma\}} \mathcal{O}(\sigma) \exp(-\beta H) \ , \ Z_N = \sum_{\{\sigma\}} \exp(-\beta H)$$

We will denote by $\mathbb{E}$ the average over all the others (quenched) random variables, and the free energy f_N per site and its thermodynamic limit are defined by

$$-\beta f_N = \frac{1}{N} \mathbb{E} \ln Z_N \ , \ f = \lim_{N\to\infty} f_N$$

We will use the notation Ω for the product of the needed number of independent copies (replicas) of ω and $\langle \cdot \rangle$ for the composition of an $\mathbb{E}$-type average over some quenched variables and some sort of Boltzmann-Gibbs average over the spin variables, to be specified each time. The multi-overlaps are defined (using

replicas) by

$$q_{1\cdots n} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^{(1)}\cdots\sigma_i^{(n)} \ .$$

Definition 1 *A* Random Multi-Overlap Structure $\mathcal{R}$ *is a triple* $(\Sigma, \{\tilde{Q}_n\}, \xi)$ *where*

- Σ *is a discrete space;*
- $\xi : \Sigma \to \mathbb{R}_+$ *is a system of random weights;*
- $\tilde{q}_{r_1\cdots r_{2l}} : \Sigma^{2l} \to [0,1], l \in \mathbb{N}, |\tilde{q}| \leq 1$ *is a positive definite* Multi-Overlap Kernel.

3 The Structure of the model

In order to understand what is the underlying structure of the model, it is well known that it is useful to compute the derivative of the free energy with respect to the somewhat basic parameter. In the case of non-diluted spin glasses such parameter the strength of the couplings (and it is equivalent to differentiating with respect to the inverse temperature). In the case of diluted spin glasses such parameter is the connectivity.
It is very easy to show (see e.g. [1]) by pretty standard calculation that

$$\frac{d}{d\alpha}\frac{1}{N}\mathbb{E}\ln\sum_{\gamma}\xi_\gamma\exp(-\beta H) = \sum_{n>0}\frac{(-1)^{n+1}}{n}\left(\frac{e^{-\beta}-1}{2^K}\right)^n \langle(1+Q_n(q))^K\rangle \quad (1)$$

where

$$Q_{2n}(q) = \sum_{l=1}^{n}\sum_{r_1<\cdots<r_{2l}}^{1,2n} q_{r_1\cdots r_{2l}} \ , \ Q_{2n+1} = 0 \ .$$

The fundamental quantities governing the model are therefore the multi-overlap, like for diluted spin glasses ([4, 5]). That is why we use RaMOSt in this context as well. The main difference is that here the function $1+Q_n$ takes the place of the mere multi-overlaps.

As it should be clear from [6, 4, 5], we must therefore introduce also two random variables $\tilde{H}_{\cdot}(\gamma,\alpha;\tilde{J})$ and $\hat{H}(\gamma,\alpha;\hat{J})$ such that

$$\frac{d}{d\alpha}\frac{1}{N}\mathbb{E}\ln\sum_{\gamma}\xi_{\gamma}\exp(-\beta\tilde{H}_{\cdot}) = K\sum_{n>0}\frac{(-1)^{n+1}}{n}\left(\frac{e^{-\beta}-1}{2^{K-1}}\right)^{n}\langle(1+Q_n(\tilde{q}))^{K-1}\rangle$$

$$\frac{d}{d\alpha}\frac{1}{N}\mathbb{E}\ln\sum_{\gamma}\xi_{\gamma}\exp(-\beta\hat{H}) = (K-1)\sum_{n>0}\frac{(-1)^{n+1}}{n}\left(\frac{e^{-\beta}-1}{2^{K}}\right)^{n}\langle(1+Q_n(\tilde{q}))^{K}\rangle$$

Finally, we introduce as expected the following trial function, that we write this first time only considering an external field

$$G_N(\mathcal{R},\tilde{H},\hat{H}) = \frac{1}{N}\mathbb{E}\ln\frac{\sum_{\sigma,\tau}\xi_{\tau}\exp(-\beta\sum_{i=1}^{N}(\tilde{H}_i+h)\frac{1}{2}(1+J_i\sigma_i))}{\sum_{\tau}\xi_{\tau}\exp(-\beta\hat{H})}$$

where $\tilde{H}_i$ are independent copies of $\tilde{H}_{\cdot}$ We will construct explicitly $\tilde{H}_{\cdot}$ and $\hat{H}$ in the next sections. Let us define

$$\tilde{H} = \sum_{i=1}^{N}\tilde{H}_i\frac{1}{2}(1+J_i\sigma_i)\ .$$

4 Generalized Bound and Extended Variational Principle

Let us state the extension to the K-SAT model of the results presented in [6, 4].

Consider the interpolating Hamiltonian

$$H_{\gamma}(t) = H(t) + \tilde{H}(1-t) + \hat{H}(t)$$

where t is understood to multiply the connectivity, and

$$R(t) = \frac{1}{N}\mathbb{E}\ln\frac{\sum_{\gamma,\sigma}\xi_{\gamma}\exp(-\beta H_{\gamma}(t))}{\sum_{\gamma}\xi_{\gamma}\exp(-\beta\hat{H}_{\gamma})}$$

then it is easy to prove ([4, 1, 6]) by interpolation the following

Theorem 1 (Generalized Bound)

$$-\beta f \le \lim_{N\to\infty}\inf_{\mathcal{R}} G_N\ .$$

The proof is based on observing that $R(1) = -\beta f_N, R(0) = G_N$ and computing the t-derivative of $R(t)$ using the expressions in the previous section

$$\frac{d}{dt}R(t) = -\alpha \sum_{n>0} \frac{1}{2n} \left(\frac{e^{-\beta}-1}{2^K}\right)^{2n} \times$$
$$\langle (1+Q_{2n}(q))^K - p(1+Q_{2n}(q))(1+Q_{2n}(\tilde{q}))^{K-1} + (p-1)(1+Q_{2n}(\tilde{q}))^K \rangle$$

where the odd terms are missing since they cancel out. Therefore the derivative above is non-positive since the function $x^p - pxy^{p-1} + (p-1)y^p$ of x and y is non-negative.

The Boltzmann RaMOSt $\mathcal{R}_B(M)$ is by definition the one for which $\Sigma = \{-1,1\}^M$ and, using τ instead of γ, one choses $\xi_\tau = \exp(-\beta H_M(\tau))$ and

$$\begin{aligned}
\tilde{H}_\tau &= -\sum_{\nu=1}^{P_{K\alpha N}} \frac{1}{2}(1+\tilde{J}^1_\nu \tau_{j^1_\nu}) \cdots \frac{1}{2}(1+\tilde{J}^{K-1}_\nu \tau_{j^{K-1}_\nu}) \frac{1}{2}(1+J^K_\nu \sigma_{i_\nu}) \\
\hat{H}_\tau &= -\sum_{\nu=1}^{P_{(K-1)\alpha N}} \frac{1}{2}(1+\hat{J}^1_\nu \tau_{j^1_\nu}) \cdots \frac{1}{2}(1+\hat{J}^K_\nu \tau_{j^K_\nu})
\end{aligned}$$

where the independent random variables $j^\cdot_\cdot$ are uniformly distributed over $1, \ldots, M$ and $\tilde{J}^\cdot_\cdot, \hat{J}^\cdot_\cdot$ are independent copies of J. The (limiting) Boltzmann RaMOSt $\mathcal{R}_B$ fulfills the Reversed Bound ([6, 4])

$$-\beta f \geq \lim_{N\to\infty} \liminf_{M\to\infty} G_N(\mathcal{R}_B(M)) = \lim_{N\to\infty} G_N(\mathcal{R}_B)$$

which can be proven in the same way as the analogous theorem 3 of [4], except here we must chose $\alpha' = \alpha(1+(K-1)N/M)$. As a consequence, one can state the following

Theorem 2 (Extended Variational Principle)

$$-\beta f = \lim_{N\to\infty} \inf_{\mathcal{R}} G_N \ .$$

It is easy to see that the Boltzmann RaMOSt for the K-SAT is factorized in the sense of section 5 of [4].

5 Replica Symmetry Breaking and Ultrametric RaMOSt

We are about to extend the results of [5] to the K-SAT by constructing the Ultrametric RaMOSt $\mathcal{R}_U$ with R-level Replica Symmetry Breaking. The latter corresponds to the choice of $\xi_\gamma(m_1, \dots, m_R), \gamma = (\gamma_1, \dots, \gamma_R)$ (illustrated e.g. in [2]) derived from the Random Probability Cascades, to be used in the case of the K-SAT together with

$$\begin{aligned}
\tilde{H}_\gamma &= \sum_{\nu=1}^{P_{K\alpha N}} \tilde{u}_\nu^\gamma \frac{1}{2}(1 + J_\nu \sigma_{i_\nu}) - \frac{1}{\beta} \ln\cosh(\beta \tilde{u}_\nu^\gamma) \\
\hat{H}_\gamma &= \sum_{\nu=1}^{P_{(K-1)\alpha N}} \hat{u}_\nu^\gamma - \frac{1}{\beta} \ln\cosh(\beta \hat{u}_\nu^\gamma)
\end{aligned}$$

with $\tilde{u}_\gamma, \hat{u}_\gamma$ chosen such that

$$\begin{aligned}
\tanh(\beta \tilde{u}_\gamma) &= (e^{-\beta} - 1)\frac{1}{2}(1 + \tilde{J}^1 W_\gamma^1) \cdots \frac{1}{2}(1 + \tilde{J}^{K-1} W_\gamma^{K-1}) \\
\tanh(\beta \hat{u}_\gamma) &= (e^{-\beta} - 1)\frac{1}{2}(1 + \hat{J}^1 W_\gamma^1) \cdots \frac{1}{2}(1 + \hat{J}^{K} W_\gamma^{K})
\end{aligned}$$

in which W_γ is the same as in the case of diluted p-spin glasses ([5])

$$W_\gamma = \tilde{\omega}_{\tilde{\alpha}_1}(\rho_{k_\nu})\bar{J}_{\gamma_1} + \cdots + \tilde{\omega}_{\tilde{\alpha}_R}(\rho_{k_\nu})\bar{J}_{\gamma_1 \cdots \gamma_R}$$

where $\tilde{\omega}_{\tilde{\alpha}}(\rho_{k_\nu})$ is the infinite volume limit of the Boltzmann-Gibbs average of a random spin from an auxiliary system with a Viana-Bray one-body interaction Hamiltonian at connectivity $\tilde{\alpha}$ ([5]). The indices, the bar, the tilde, the hat mean independent copies of the corresponding variables. Let us now report a comment from [4]. Given any partition $\{x^a\}_{a=0}^R$ of the interval $[0, 1]$, there exists a sequence $\{\tilde{\alpha}_a\}_{a=0}^R \in [0, \infty]$ such that $\tilde{q}_{1\cdots n}(\tilde{\alpha}_a) = x_a - x_{a-1}$. In other words, a sequence $\{\tilde{\alpha}_a\}_{a=0}^R \in [0, \infty]$ generates for each $n \in \mathbb{N}$ a partition of $[0, 1]$ considered as the set of trial values of $\tilde{q}_{1\cdots n}$, provided the $\tilde{\alpha}_a$ are not too large

$$\sum_{a \leq R} \tilde{q}_{1\cdots n}(\tilde{\alpha}_a) \leq 1 \ . \tag{2}$$

We limit our trial multi-overlaps to belong to partions generated in this way. This implies that the points of the generated partitions tend to get closer to

zero as n increases. Which is good, since in any probability space $\langle \tilde{q}_n \rangle$ decreases as n increases and therefore the probability integral distribution functions tend to grow faster near zero. Now put inductively

$$\mathbb{E}\tilde{\Omega}_{\tilde{\alpha}_a}(\tau_{k_\nu}^{(r_1)} \cdots \tau_{k_\nu}^{(r_l)}) = \tilde{q}_{r_1 \cdots r_l}(\tilde{\alpha}_a) = \tilde{q}_{r_1 \cdots r_l}^{(a)} - \tilde{q}_{r_1 \cdots r_l}^{(a-1)} \ , \ \tilde{q}_{r_1 \cdots r_l}^{(0)} = 0$$

then an elementary calculation shows that

$$\begin{aligned} \mathbb{E} \tanh^n(\beta \tilde{u}_\gamma) &= \left(\frac{e^{-\beta} - 1}{2^{K-1}} \right)^n (1 + Q_n(\tilde{q}))^{K-1} \\ \mathbb{E} \tanh^n(\beta \hat{u}_\gamma) &= \left(\frac{e^{-\beta} - 1}{2^{K}} \right)^n (1 + Q_n(\tilde{q}))^{K} \end{aligned}$$

with $\tilde{q}$ ultrametric ([5]), i.e.

$$\tilde{q}_{r_1 \cdots r_l} = (\tilde{q}_{r_1 \cdots r_l}^{(1)} - \tilde{q}_{r_1 \cdots r_l}^{(0)}) \delta_{\gamma_1^{r_1} \dots \gamma_1^{r_l}} + \cdots + (\tilde{q}_{r_1 \cdots r_l}^{(R)} - \tilde{q}_{r_1 \cdots r_l}^{(R-1)}) \delta_{\gamma_1^{r_1} \dots \gamma_1^{r_l}} \cdots \delta_{\gamma_R^{r_1} \dots \gamma_R^{r_l}}$$

If we denote by X the map $\tilde{\alpha}_a \to m_a$ satisfying (2), we have $\mathcal{R}_U = \mathcal{R}_{X,R}$ and the trial function is

$$\inf_{X,R} G(\mathcal{R}_{X,R}) \ .$$

Notice that the Ultrametric RaMOSt for the K-SAT factorizes exactly like the Boltzmann one, which is optimal.

6 Conclusions

The RaMOSt is the minimal generalization of the ROSt, and what we showed here and in [4] is that the minimal generalization is enough to formulate the variational principle and also exhibit a concrete RaMOSt analogous to the Parisi one for SK. As a consequence, it is enough to restrict the space of trial functions to those expressible in terms of fixed multi-overlaps (i.e. a set of numbers, not random variables to be averaged).

A The p-XOR-SAT

The Hamiltonian of the random p-XOR-SAT coincides with the one of the diluted p-spin glass

$$H = -\sum_{\nu=1}^{P_{\alpha N}} J_\nu \sigma_{i_\nu^1} \cdots \sigma_{i_\nu^p}$$

It is therefore elementary to extend all the results of [4, 5] to this case, also when in presence of an external field. Since it is easy to show ([4])

$$\frac{d}{d\alpha}\frac{1}{N}\mathbb{E}\ln\sum_{\gamma}\xi_{\gamma}\exp(-\beta H)=\sum_{n>0}\frac{1}{2n}\mathbb{E}\tanh^{2n}(\beta J)(1-\langle q_{2n}^{p}\rangle)$$

the structure of the model is the same RaMOSt valid for the case of the Viana-Bray model, but the equality above suggests to try and get the non-negative convex function $x^p-pxy^{p-1}+(p-1)y^p$ whenever we got the square $x^2-2xy+y^2$ in the Viana-Bray case.

That is why here $\tilde{H}$ and $\hat{H}$ are chosen such that

$$\begin{aligned}\frac{d}{d\alpha}\frac{1}{N}\mathbb{E}\ln\sum_{\gamma}\xi_{\gamma}\exp(-\beta\tilde{H}.) &= p\sum_{n>0}\frac{1}{2n}\mathbb{E}\tanh^{2n}(\beta J)(1-\langle\tilde{q}_{2n}^{p-1}\rangle)\\ \frac{d}{d\alpha}\frac{1}{N}\mathbb{E}\ln\sum_{\gamma}\xi_{\gamma}\exp(-\beta\hat{H}) &= (p-1)\sum_{n>0}\frac{1}{2n}\mathbb{E}\tanh^{2n}(\beta J)(1-\langle\tilde{q}_{2n}^{p}\rangle)\end{aligned}$$

and plugged in

$$G_N(\mathcal{R},\tilde{H},\hat{H})=\frac{1}{N}\mathbb{E}\ln\frac{\sum_{\sigma,\tau}\xi_{\tau}\exp(-\beta\sum_{i=1}^{N}(\tilde{H}_i+h)\sigma_i)}{\sum_{\tau}\xi_{\tau}\exp(-\beta\hat{H})}$$

The Generalized Bound clearly holds, the Boltzmann RaMOSt is the one with

$$\tilde{H}_{\tau}=-\sum_{\nu=1}^{P_{p\alpha N}}\tilde{J}_{\nu}\tau_{j_{\nu}^{1}}\cdots\tau_{j_{\nu}^{p-1}}\sigma_{i_{\nu}}\ ,\ \hat{H}_{\tau}=-\sum_{\nu=1}^{P_{(p-1)\alpha N}}\hat{J}_{\nu}\tau_{j_{\nu}^{1}}\cdots\tau_{j_{\nu}^{p}}\ ,\ \xi_{\tau}=e^{-\beta H_M}$$

(same couplings as the original system) and it is optimal so that we can also state the Extended Variational Principle. The Broken Replica Symmetry Ultrametric RaMOSt (which includes as a trivial case the Replica Symmetric one) relies on the weights ξ_{γ} of the Random Probability Cascades as in [2] and on

$$\begin{aligned}\tilde{H}_{\gamma} &= -\sum_{\nu=1}^{P_{p\alpha N}}\left(\frac{1}{\beta}\ln\frac{\cosh(\beta J)}{\cosh(\beta\tilde{J}_{\nu}^{\gamma})}+\tilde{J}_{\nu}^{\gamma}\sigma_{i_{\nu}}\right)\\ \hat{H}_{\gamma} &= -\sum_{\nu=1}^{P_{(p-1)\alpha N}}\left(\frac{1}{\beta}\ln\frac{\cosh(\beta J)}{\cosh(\beta\hat{J}_{\nu}^{\gamma})}+\hat{J}_{\nu}^{\gamma}\right)\end{aligned}$$

with

$$\tanh(\beta\tilde{J}_{\gamma})=\tanh(\beta J)\tilde{W}_{\gamma}^{1}\cdots\tilde{W}_{\gamma}^{p-1}\ ,\ \tanh(\beta\hat{J}_{\gamma})=\tanh(\beta J)\tilde{W}_{\gamma}^{1}\cdots\tilde{W}_{\gamma}^{p}$$

where $\tilde{W}_{\gamma}^{\cdot}$ are independet copies of

$$\tilde{W}_{\gamma}(\bar{J}, k_{\nu}) = \tilde{\omega}_{\tilde{\alpha}_1}(\rho_{k_{\nu}})\bar{J}_{\gamma_1} + \cdots + \tilde{\omega}_{\tilde{\alpha}_R}(\rho_{k_{\nu}})\bar{J}_{\gamma_1 \cdots \gamma_R}$$

with $\bar{J}_{\cdot} = \pm 1$ symmetric.

Acknoledgments

The author warmly thanks Francesco Guerra for useful discussions, and gratefully acknowledges the hospitality of the Department of Physics at University of Rome "La Sapienza" (and in particular Giovanni Jona-Lasinio).

References

[1] S. Franz, M. Leone, *Replica bounds for optimization problems and diluted spin systems*, J. Stat. Phys. **111** 535 (2003)

[2] D. Panchenko, M. Talagrand, *Bounds for diluted mean-field spin glass models*, to appear in Prob. Theory Related Fields

[3] M. Mezard, G. Parisi, *The Bethe lattice spin glass revisited*, Eur. Phys. J. B **20**, 217 (2001)

[4] L. De Sanctis, *Random Multi-Overlap Structures and Cavity Fields in Diluted Spin Glasses*, ArXiv:cond-mat/0403506

[5] L. De Sanctis, *Ultrametric Broken Replica Symmetry RaMOSt*, ArXiv:cond-mat/

[6] M. Aizenman, R. Sims, S. L. Starr, *An Extended Variational Principle for the SK Spin-Glass Model*, Phys. Rev. B, **68**, 214403 (2003)